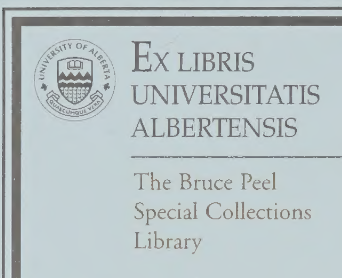

UNIVERSITY OF ALBERTA
QUAECUMQUE VERA

University of Alberta

Library Release Form

Name of Author: Francis Blais

Title of Thesis: Le Canadian Achievement Test utilisé comme un indicateur de réussite scolaire.

Degree: Master of Education

Year this Degree Granted: 2001

University of Alberta

Le Canadian Achievement Test utilisé comme un indicateur de réussite scolaire.

par

Francis Blais 

Mémoire présenté à la Faculty of Graduate Studies and Research en vue de l'obtention du diplôme de Maîtrise en sciences de l'éducation - Etudes en langue et culture

Faculté Saint-Jean

Edmonton, Alberta

Printemps 2001

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled “Le Canadian Achievement Test utilisé comme un indicateur de réussite scolaire.” submitted by Francis Blais in partial fulfillment of the requirements for the degree of Master of Education in Educational Studies in Language and Culture.

Résumé

Le but de cette étude est d'établir s'il y a des corrélations significatives entre les résultats du Canadian Achievement Test (CAT) et la note finale attribuée par l'école à des étudiants de neuvième année en anglais et en mathématiques.

L'hypothèse de base est qu'il y aura une tendance significative qui émergera des résultats de cette recherche et que cette tendance pourra être utilisée comme un indicateur de prédiction de la réussite scolaire d'un élève en anglais et mathématique. La population de cette étude est composée de120 étudiants de neuvième année d'une école secondaire premier cycle. De cette population, 78 étudiants proviennent du programme d'anglais régulier, 18 étudiants du programme d'éducation spéciale en anglais et 24 étudiants du programme d'immersion française.

Les sections lecture, langage, habiletés d'étude et mathématique du CAT pour toute l'échantillon ont obtenues des coefficients de corrélation significatifs en fonction de la note finale obtenue par des étudiants en anglais et en mathématique. Ce qui veut dire qu'on peut utiliser ces parties du CAT comme des indicateurs de réussite scolaire. Par contre les calculs de corrélation obtenus pour la section épellation et la note finale obtenue par les étudiants en anglais ne sont aucunement significatif et ne peuvent pas être utilisés comme un indicateur de réussite scolaire.

Abstract

The purpose of this research was to examine whether there is a significant correlation between the Canadian Achievement Test (CAT) results and the final school scores given to grade nine students in English and Mathematics. The main hypothesis of this research is that a significant relationship will be found between the CAT results and the final grade nine scores and that this relationship could be used as an indicator to predict the success of students in English and Mathematics. The sample is composed of one hundred and twenty grade nine students from one school who wrote the CAT and had final school-based grades. Of these students, seventy-eight were enrolled in the regular English program, eighteen were enrolled in the Special Education program and twenty-four were enrolled in the French Immersion program.

The results of this study found significant correlations between language, reading, study skills and mathematics subscores of the CAT and final grades in English and Mathematics. This indicates that these subscores of the CAT may be used as indicators of school success in English and Mathematics. However, the subscore in the spelling section of the CAT was not found to be significantly correlated with final grades in English and Mathematics and therefore would not be useful as an indicator of school success in these subjects.

Table des Matières

Liste des Tableaux

CHAPITRE I

Introduction

Description du contexte

La situation dans le monde de l'éducation en Alberta en 1999 a changé beaucoup depuis les dix dernières années. L'apparition des tests de rendement scolaire provincial (Provincial Achievement Tests) administrés aux étudiants de troisième, sixième et neuvième année dans les sujets majeurs (anglais, français, mathématique, etc.) a catalysé un mouvement de changements et d'inquiétude dans le domaine de l'éducation comme le mentionne Runté (1998). Runté mentionne aussi que certains journaux locaux publient les résultats de ces tests de rendement et que des comparaisons sont faite entre les écoles et les commissions scolaires sans tenir compte de plusieurs facteurs tels que les facteurs socio-économiques. Il mentionne aussi que les enseignants se plaignent du fait que certains administrateurs se basent sur ces résultats de tests pour les évaluer au niveau de leur compétence à enseigner. Plusieurs problématiques qui peuvent être adressées de différentes manières. En plus, un nombre d'enseignants et de parents restent sceptiques face aux résultats de tests standardisés comme le Canadian Achievement Test qui ne sont pas mandatés par la province et qu'on utilisent dans certaines commissions scolaires comme à St-Albert ou à Edmonton.

Impara et Plake (1998) décrivent le Canadian Achievement Test, communément appelé le CAT, comme une batterie de tests qui a été créé en 1981 par le personnel de la compagnie CTC/McGraw-Hill Ryerson pour mesurer le rendement des étudiants aux niveaux des habiletés de base en mathématique et dans les arts du langage en anglais. Ils mentionnent que le CAT est un test national qui a été créé pour mesurer le rendement des étudiants de la deuxième à la douzième année. Même si ce test n'est pas relié à un programme d'enseignement particulier, il a été construit pour diagnostiquer les forces et les faiblesses des étudiants vis-à-vis les différents programmes d'études. Ces tests



soulèvent aussi plusieurs questions. Est-ce que les résultats de ces tests standardisés peuvent être utilisés comme un outil d'évaluation concret ou plutôt comme un indicateur de réussite scolaire des étudiants? On définit par réussite scolaire qu'un étudiant reçoit une note de 50% ou plus dans une matière. Qu'en est-il de la note finale donnée par une école que reçoit un étudiant à la fin de l'année dans une matière? Est-ce que cette note a une certaine validité au niveau de l'évaluation ou est-ce simplement un indicateur de réussite scolaire de l'élève dans une matière? Jusqu'à maintenant, aucune recherche n'a été faite sur la corrélation des résultats du Canadian Achievement Test et la note finale en mathématique et en anglais.

L'objectif principal de ce travail est d'établir s'il y a une corrélation positive significative entre les résultats du Canadian Achievement Test et la note finale en mathématique et en anglais données par l'école aux étudiants en neuvième année. Les calculs de corrélation entrés serviront de base comme un indicateur de réussite scolaire d'un éleve de la neuvième année soit en anglais régulier, anglais éducation spéciale ou immersion française. Si les calculs de corrélation s'avèrent d'une tendance positive significative, les professeurs pourront s'en servir comme un outil de prédiction de réussite scolaire en anglais et en mathématique donc comme un outil diagnostique.

But de l'étude

Le but de cette étude est d'établir s'il y a des corrélations positives significatives entre les résultats du Canadian Achievement Test (CAT) et la note finale attribuée par l'école à des étudiants de neuvième année en anglais et en mathématiques. S'il y a une corrélation significative positive, on pourra se servir des résultats du CAT comme un outil de prédiction de la réusssite scolaire d'un étudiant en anglais et en mathématique donc comme un outil diagnostique.

La note finale en anglais et en mathématique des étudiants de cette recherche est obtenue par la compilation de notes données pour des tests, travaux, quiz, etc, par les

professeurs durant l'année scolaire qui correspond à 80% de la note finale et la note reçue lors de l'examen de rendement de la province écrit en juin qui correspond à 20% de la note finale de l'étudiant.

L'hypothèse de base est qu'il y aura une tendance significative qui émergera des résultats de cette recherche et que cette tendance pourra être utilisée comme un indicateur de prédiction de la réussite scolaire d'un élève en anglais et mathématique donc comme un outil diagnostique. La population de cette étude est composée de120 étudiants de neuvième année dont 78 dans le programme d'anglais régulier, 18 dans le programme d'éducation spéciale en anglais et 24 dans le programme d'immersion française. Toutes les notes finales et les résultats du CAT ont par la suite été transférés par le chercheur dans le logiciel Excel et les calculs de corrélations ont été faits. Puisqu'aucune étude n'a été faite au sujet de la corrélation entre les résultats du Canadian Achievement Test et la note finale en anglais et en mathématique les corrélations concernant ce même test et la note finale d'un étudiant en mathématique et en anglais, cette étude démontrera son utilité. Elle sera utile pour tout professeur, école ou district scolaire voulant utiliser un indicateur de prédiction de la réussite scolaire d'un étudiant en plus d'assister le milieu scolaire à apporter l'aide nécessaire à un étudiant éprouvant des difficultés en mathématique et en anglais.

CHAPITRE II

Revue de la Littérature

L'utilisation de tests standardisés depuis plusieurs années dans le monde de l'éducation a été très fréquente. On a tenté d'utiliser les résultats de tests standardisés à plusieurs fins tels que dans les milieux politiques pour changer les programmes en éducation par exemple, et économique pour justifier des coupures budgétaires ou l'octroi de subventions en éducation. Dans le milieu scolaire ces tests ont été utilisés comme

outil d'évaluation ou agent de prédiction de succès académique et pour établir des coefficients de corrélation entre le rendement scolaire des étudiants et les résultats de tests standardisés.

Premièrement au niveau économique et politique Reis, Westberg, Kulikowich et Purcell (1998) utilisent les résultats de tests standardisés en relation avec la compression des programmes en éducation causée par les coupures budgétaires des gouvernements. Baker (1988) discute que les tests standardisés ont changé leur fonction primaire à travers les années. Au début, ils devaient être des indicateurs que le monde de l'éducation accomplissait son but d'éduquer les élèves. Maintenant ils sont des outils de politiques éducationnels qui servent comme preuves du succès des décisions politiques prises par les gouvernements. Ils sont aussi devenus des moyens efficaces de changer la productivité dans les écoles en les comparant au rendement éducationnel c'est-à-dire des notes plus élevées obtenues par les étudiants.

McCurdy et Speich (1991) approchent les tests standardisés aux Etats-unis sous différentes facettes. Entre autres, ils décrivent l'historique des tests standardisés et le mouvement de recul vis-à-vis les tests standardisés dans la réforme scolaire des années 1990. De plus, ils traitent de la question de remplacer les tests standardisés à choix multiples par les tests qui sont créés par les enseignants.

Deuxièmement, les tests standardisés ont été utilisés par certains éducateurs comme un outil de placement ou d'évaluation de succès scolaire. Gronna, Jenkins, Chin-Chance et Selvin (1998) se servent des tests standardisés pour analyser les performances académiques des étudiants avec handicap dans un programme de tests standardisés. Ils ont comparé les données d'une partie du Stanford 8 Test administré à des enfants avec différents handicaps soit au niveau du langage, de l'apprentissage, de l'état émotionnel ou d'une légère arriération mentale et la batterie complète du Stanford 8 avec une population sans handicap dite normative. Lorsqu'analysés, les résultats ont procuré que des comparaisons limitées entre la population d'élèves handicapés et la

population sans handicap.

Ascher (1990) a utilisé les résultats de tests de rendement standardisés et de tests d'aptitudes pour prouver qu'ils n'avaient qu'une valeur très limitée pour le placement académique d'étudiants bilingues provenant de milieu hispanique ou asiatique. Ascher a découvert que les résultats de ces deux tests favorisaient les étudiants unilingues anglais car les tests étaient biaisés pour les étudiants dont l'anglais n'était pas leur langue maternelle et on interpréte les résultats inférieurs des étudiants bilingues comme étant des preuves qu'ils avaient des difficultés d'apprentissages ou autres. Les options proposées par Ascher pour éliminer les biais dans les tests sont les suivantes: utiliser des tests non-verbaux, traduire les tests, utiliser des interprètes, utiliser des tests qui réfèrent aux normes de la langue maternelle.

Troisièmement, certains tests se servent de tests standardisés comme un outil de prédiction. Une recherche faite par Hafner (1990) en Floride, a essayé d'établir certaines variables de prédiction en utilisant les résultats d'un test standardisé, le California Achievement Test. Les résultats utilisés sont pour le vocabulaire, la compréhension en lecture, le langage, la mathématique, la science et en études sociales et les résultats de plusieurs tests standardisés tel que le vocabulaire, l'acquisition d'apprentissages, l'arithmétique, la reconnaissance de mots et les habiletés d'écoute. Les résultats de sa recherche ont permis d'établir des corrélations variant entre un minimum de 0.77 pour la compréhension en lecture et le vocabulaire et un maximum de 0.88 pour la mathématique et ces résultats ont amené des chercheurs à faire des suggestions, basées sur les résultats de Hafner, pour améliorer le rendement des étudiants au niveau des habiletés d'écoute, des habiletés de reconnaissance de mots et au niveau de la quantité et de la qualité d'emmagasiner du vocabulaire et de l'information générale utilisée dans la vie courante. Quelques suggestions pour améliorer le vocabulaire, les habiletés d'écoute, la reconnaissance de mots, sont que les parents doivent lire à leurs enfants dès le plus jeune âge, que les enfants doivent être entourés de livres, que l'école doit utiliser une diversité

de matériel didactique. Il mentionne qu'il faut des activités spécifiques d'écoute et de structure pour améliorer les habiletés d'écoute des enfants et qu'il faut des unités d'enseignement spécifique en géographie, histoire, langage, pour ne nommer que ceux-ci, pour améliorer le niveau de compréhension en lecture.

Beard (1990) s'est servi des résultats de tests standardisés utilisés par un district scolaire de la Floride, le Norm-Referenced Test (NRT) pour prédire la performance des étudiants en troisième, cinquième, huitième et dixième année sur un test employé par l'état de la Floride intitulé State Student Assesment Test (SSAT). Les résultats de cette étude ont démontré que les résultats du NRT peuvent être utilisés pour exempter des étudiants d'écrire le SSAT. Ils ont aussi démontré que 75% des étudiants ayant écrit le NRT ont eu une performance non satisfaisante sur le SSAT. Johnson (1989), dans son étude ayant trait au processus d'apprentissage et aux différentes variables en lecture, voulait savoir si les variables au niveau du processus d'apprentissange et en lecture d'élèves en première année universitaire pourraient servir d'indicateur de prédiction de leur note moyenne en première année d'université. Johnson a pris certaines variables au niveau du processus d'apprentissage et en lecture. Elle a pris ces variables et les a combinées avec des résultats de tests standardisés comme le American College Testing et le Scholastic Aptitude Test pour voir si elle pouvait garder leur valeur indépendante de prédiction. Son étude a démontré que les variables en lecture et au niveau du processus d'apprentissage établissait un indicateur de prédiction significatif de leur note moyenne en première année d'université.

En dernier lieu, on utilise des calculs de corrélation entre des notes finales de semestre d'étudiants de niveau secondaire et des résultats de tests standardisés et entre des résultats de deux tests standardisés. Olson (1989) a tenté d'établir s'il y avait des corrélations entre la note du test final du semestre, la note finale obtenue par l'étudiant et les résultats d'un test standardisé appelé Survey Tests of Essential Elements/Learner Standards. Sur les 44 tests qu'il a examinés à tous les niveaux du secondaire, il a trouvé



qu'il n'y avait aucune tendance significative pouvant être établie entre le test standardisé, la note du test final de semestre et la note finale de l'étudiant. Quand Perlman (1988) a administré à des élèves de troisième, sixième et huitième années deux tests standardisés, le Illinois Goal Assesment Program et le Iowa Tests of Basic Skills mesurant le niveau de signification littéral, de lecture, d'inférence et de généralisation, elle a trouvé qu'il y avait une corrélation forte entre le niveau de signification littéral du ITBS et le niveau d'inférence de compréhension du IGAP.

En résumé, les résultats de tests standardisés ont été utilisés par les chercheurs de plusieurs façons. Soit au niveau politique et économique pour justifier la réduction des budgets scolaires par les gouvernements, causée par leurs coupures budgétaire. Deuxièmement pour évaluer indirectement la productivité dans les écoles en comparant les montants d'argent attribués aux écoles par le gouvernement au rendement éducationnel c'est-à-dire des notes plus élevées obtenues par les étudiants. Troisièmement pour remettre en question l'utilisation des tests standardisés ou comme un outil d'évaluation ou de classement scolaire ou comme un indicateur de prédiction de la réussite scolaire. En dernier lieu, pour le calcul de corrélation pour établir s'il y a des relations significatives entre des résultats de tests standardisés et d'autres variables dans le domaine de l'éducation comme la note finale attribuée à un étudiant en anglais et en mathématique.

Historique des tests standardisés

Cette section porte sur une vue historique et génerale des tests standardisés et des tests de rendement sur une période de temps assez étendue de 1900 jusqu'à aujourd'hui. Par la suite, l'utilisation des tests standardisés pour mesurer le rendement; leurs avantages et leurs inconvénients sera discuté ainsi qu'une description du Canadian Achievement Test et de son utilisation dans le milieu scolaire.

L'utilisation des statistiques et des tests standardisés ne datent pas d'hier.

Depuis les années 1900, les tests standardisés et de rendement ont été utilisés de plus en plus. Mehrens et Lehmann (1991) mentionnent que la raison pour utiliser les tests standardardisés est d'obtenir des échantillons de comportements en utilisant des procédures uniformes.

Dans le début des années 1900, le monde de la psychologie et de l'éducation aux Etats-unis utilisaient déjà les transformations linéaires (Standard Scores) pour fins d'analyse statistiques. Angoff (1984) mentionne que durant cette période on se servait de certains tests et méthodes d'analyses statistiques. Par exemple, Angoff mentionne qu'en 1922 Hull utilisait une des plus simples méthodes de mesure telle que la transformation linéaire non-ajustée. Ce type de test était administré à un groupe d'individus considérés en quelque sorte comme un groupe standard de référence. Dans d'autres occasions, les individus étaient tirés au hasard d'une population définie avec certaines caractéristiques définies. Angoff raconte que les individus représentaient parfois un groupe pré-défini, et ils étaient considérés d'être similaires en fonction de la population à laquelle le test était dirigé.

Durant les années 1930 à 1960 Angoff mentionnent que plusieurs chercheurs de renommée internationale ont eux aussi utilisés les résultats standards normalisés de tests pour fins de recherches. Angoff mentionne que McCall en 1939 a été principalement associé avec les résultats standards normalisés. Il mentionne aussi que d'autres chercheurs tels que Pearson en 1906, E. L. Thorndike en 1927 et Thurstone en 1925 sont aussi connus pour l'utilisation de résultats standards normalisés. Les termes rang percentile, distribution, échantillon, population, échelle de pointage brut (Raw Score Scale), sont aussi apparus de plus en plus fréquemment durant cette période aidant par le fait même à établir une terminologie propre aux tests standardisés.

Après la deuxième guerre mondiale, la plus grande utilisation des tests standardisés et de leurs résultats n'ont pas qu'introduit que validité et fiabilité; la validité est l'assurance qu'un test mesure ce qu'il doit mesurer et rien d'autre et la

fidélité est l'assurance qu'un test mesure toujours de la même façon. Les chercheurs ont aussi apporté des questions critiques sur ceux-ci. Ebel et Frisbie (1986) mentionnent plusieurs cas aux Etats-unis dans lesquels les tests, ou les résultats des tests, ont été contestés pour diverses raisons telle que la discrimination raciale et culturelle, leur manque de relation pour prédire un futur emploi et ils ont mêmes été utilisés par un étudiant pour une poursuite légale contre une institution scolaire.

Entre 1960 et 1980, quelques critiques à l'égard des résultats des tests standardisés ont refait surface dans le système légal. Ebel et Frisbie mentionnent le cas de Stell contre Savannah en 1964. Les avocats de ces deux étudiants de race blanche ont argumenté que la déségrégation des écoles du conseil scolaire de Savannah en Georgie serait nuisible pour tous les enfants à cause des différences de résultats dans les habiletés testées entre les étudiants de race blanche et noire. La cour civile de Georgie a accepté cette version mais la cour d'appel l'a rejetée et a ordonné au conseil scolaire de Savannah de continuer la déségrégation.

Ebel et Frisbie parlent aussi du cas de Bakke contre la Californie en 1978 qui parle de la discrimination raciale et de résultats de test standardisé. Alan Bakke qui est de race noire s'est vu refusé son admission à l'école de médecine de Davis à l'université de Californie. Il a découvert que d'autres étudiants de races différentes à la sienne avec des résultats de tests inférieurs aux siens ont été acceptés. Il intenta des poursuites légales contre l'université. La court suprême jugea que l'université de Californie devrait admettre Bakke et que l'utilisation de différents standards pour des étudiants de différentes races était inappropriée. Cependant, la cour a aussi suggéré qu'il était approprié que la race d'un étudiant soit considérée lorsque des décisions au niveau des admissions sont faites.

En 1971, les résultats de tests standardisés ont fait manchette, cette fois-ci dans le domaine du travail. Ebel et Frisbie mentionnent le cas de Griggs contre la compagnie Duke Power aux Etats-unis en 1971. Griggs s'était vu refusé un emploi avec cette

compagnie à cause de ses résultats de diplôme de fin d'études secondaires et de notes de passage de tests standardisés. La cour jugea que la compagnie ne pouvait pas se servir des résultats de tests standardisés à moins que ceux-ci aient des relations spécifiques avec les critères de succès pour un emploi.

A travers des années l'utilisation des tests standardisés et de leurs résultats dans plusieurs domaines en éducation a augmenté de manière continuelle et significative. Perrone (1991) mentionne que, depuis la publication en 1983 de A Nation at Risk aux Etats-unis l'utilisation des tests standardisés dans les écoles élémentaires et secondaires a augmentés beaucoup et qu'on les utilise à plusieurs fins telle que le placement d'étudiants dans les programmes avancés ou surdoués. Il discute aussi dans son article que les enseignants se sentent obligés de préparer leurs étudiants à écrire ces tests standardisés et que peu d'enseignants croient aux résultats de ces tests. Il ajoute que les enseignants ne croient pas que les résultats de tests standardisés peuvent représenter de façon précise, par exemple le niveau d'intelligence d'un étudiant. D'autres raisons énumérées par Perrone contre l'utilisation des tests standardisés sont: la possibilité de la perte d'estime de soi de l'étudiant, la distortion des programmes et la baisse des attentes des étudiants.

Les tests standardisés en langues secondes sont utilisés aux Etats-unis pour plusieurs raisons. On les utilise soit pour mesurer le niveau du langage des étudiants souvent l'anglais ou d'autres langues comme un outil pour prédire leur réussite à l'université en fonction de certains critères comme la race, le sexe et leur moyenne académique à la fin du secondaire. Rosa (1989) compare les résultats de deux tests standardisés: La Prueba Riverside de Realizacion en Espanol (PRRE) et le Spanish Assessment of Basic Education (SABE) qui mesurent le niveau de la lecture et de la mathématique des étudiants de la première à la huitième année. Ces tests ont été traduits dans la langue maternelle des étudiants, c'est-à-dire l'espagnol. La population comprenait 2634 étudiants inscrits dans 19 écoles élémentaires de Chicago qui



participaient à un programme d'éducation bilingue en espagnol. Les instruments de mesure incluaient des sous-tests en lecture et mathématique. Les tests ont été administrés par des enseignants bilingues ayant reçus un entraînement en espagnol dans le but de les préparer à participer dans l'étude. Les résultats ont indiqué que le PRRE (La Prueba Riverside de Realizacion en Espanol) et le SABE (Spanish Assesment of Basic Test) étaient acceptables pour la population évaluée. Ce qui veut dire que les deux tests sont essentiellement similaires au niveau des propriétés psychométriques dans les caractéristiques des tests classés par les enseignants.

Sue et Abe (1984) utilisent des données provenant de diverses sources telle que la moyenne académique à la fin du secondaire et les résultats de tests standardisés comme un type d'indicateur de réussite à l'université. Cette étude a examiné les indicateurs de réussite académique de 4113 américains d'origine asiatique et de 1000 étudiants de race blanche inscrits en première année universitaire dans les huit campus de l'université de Californie durant l'automne 1984. Plus spécifiquement, Sue et Abe ont mesuré: (1) la moyenne académique au secondaire; (2) le pointage de la section verbale du Scholastic Aptitude Test (SAT); (3) le pointage de la section mathématique du même test (SAT); (4) le pointage du test de composition en anglais; et (5) le pointage du test en mathématique. Ces indicateurs ont été mesurés en fonction des variables suivantes: (1) la race; (2) le champs d'étude principal; (3) la langue parlée; et (4) le genre. Les deux résultats majeurs de la recherche furent les suivants: la moyenne académique au secondaire et les résultats du test standardisé SAT peuvent prédire de façon modérée la note des étudiants asiatiques et blancs de première année universitaire; le meilleur indicateur de réussite pour les deux groupes est la moyenne académique à la fin du secondaire.

Miller et Pennington (1988) ont quant à eux essayé d'établir des coefficients de corrélation entre les résultats de deux tests standardisés, le Degrees of Reading Power Test et le Stanford Achievement Test, et de déterminer comment ces deux instruments

de mesure pourraient prédire le rendement académique des étudiants de niveau secondaire premier cycle. Les résultats des tests standardisés ont par la suite été corrélés avec la note finale du premier semestre en science, en anglais et en études sociales. Les résultats ont démontré une forte corrélation entre les deux tests mais Miller et Pennington ont conclu qu'ils étaient des indicateurs de niveau faible à modéré pour prédire le rendement des étudiants et que les enseignants devraient utiliser plus que des calculs de corrélation pour mesurer le rendement d'un étudiant ou pour prédire la performance d'une classe.

Pratt (1987) a essayé d'établir un genre d'indicateur de réussite en fonction d'un test standardisé, le National Teachers Examination en Caroline du Nord aux Etats-unis et la note de cours offerte à l'université à de futurs enseignants en utilisant des calculs de coefficient de corrélation. En examinant les résultats du test et la note des quatre cours pré-requis des étudiants, un ensemble de données pour 64 étudiants a été identifié et des analyses de régression multiple ont été calculées pour déterminer la meilleure combinaison de variables pour prédire les résultats du test. La formule fut alors appliquée à d'autres étudiants en éducation par Pratt et une très forte corrélation fut découverte entre les résultats actuels du test et les résultats majeurs du test standardisé.

En dernier lieu, une étude faite par le Canadian Education Statistics Council à Edmonton en Alberta (1994) avait pour but d'identifier les provinces et terrritoires canadiens qui employaient des tests standardisés. Et de déterminer s'ils étaient mandatés par leur gouvernement de s'en servir et s'il y avait un test standardisé commun utilisé par toutes les provinces et territoires canadiens. Le Canadian Education Statistics Council a trouvé qu'aucun instrument semblable de mesure standardisé est utilisé par toutes les provinces canadiennes et que les deux tests standardisés employés le plus fréquemment sont le Canadian Achievement Test (CAT) et le Canadian Test of Basic Skills (CTBS). Seulement Terre-Neuve et le Yukon ont mandanté de façon officiel

l'emploi de tests standardisés tandis quel'Ile-du-Prince-Edouard n'avait aucun type d'évaluation provinciale en 1993-94.

En Alberta au début des années 1990 la province a mandaté l'utilisation des tests de rendement provinciaux (Provincial Achievement Test) pour les niveaux de la troisième, sixième et neuvième année dans les sujets suivants: le français (pour les étudiants francophones et de programmes d'immersion française), l'anglais, la mathématique, la science et les études sociales. Au secondaire deuxième cycle, les examens de la province (Diploma Tests) sont administrés pour les sujets suivants: l'anglais, les études sociales, les mathématiques, la biologie, la chimie, la physique et le français pour les écoles françaises et d'immersion française. (Alberta Department of Education, 1987).

Ces tests ne sont pas considérés comme étant des tests standardisés et la province de l'Alberta n'oblige pas les commissions scolaires d'utiliser un ou plusieurs tests standardisés jusqu'à ce jour. (Student's Evaluation Branch, Alberta Education, February 2000). Présentement en l'an 2000, plusieurs écoles en Alberta utilisent le Canadian Achievement test même si l'administration de ce test n'est pas mandatée par la province.

Utilisation des Tests Standardisés pour Mesurer le Rendement Académique

Les questions qu'on doit se poser maintenant sont les suivantes: Quelles sont les caractéristiques ou les mesures utilisées dans les tests standardisés? Est-ce qu'il y a des problèmes reliés aux tests standardisés? Quels sont les avantages et les inconvénients des tests standardisés?

Caractéristiques et mesures utilisées dans les tests standardisés

Les échelles de mesures utilisées dans les tests standardisés peuvent varier d'un

test à l'autre mais on retrouve souvent des mesures qui sont similaires dans la plupart de ces tests comme les échelles de pointage brut (raw score scales), le pourcentage de réponses correctes (percent of correct answer), les rangs percentiles (percentile ranks) et le niveau d'équivalence (grade equivalents) pour ne nommer que ceux-ci. Green (1986) discute des mesures utilisées dans les tests de rendement standardisés et identifie les mesures qui sont plus ou moins efficaces. Les mesures désirées pour les tests standardisés sont les suivantes: les échelles de pointage brut et de pointage, le pourcentage de réponses correctement répondues, les rangs percentiles, le niveau d'équivalence et les mesures de courbe normale.

Green mentionne aussi que les mesures de courbe normale sont aussi bonnes que n'importe quelle mesure en autant qu'on soit confortable avec ce type de mesure. Les échelles de pointage sont les plus efficaces pour repérer des suites de données chez un ou plusieurs individus. Les échelles de pourcentage de réponses correctes sont utiles pour évaluer la quantité de connaissances ou le degré de maîtrise de ces connaissances. Par contre, ces échelles de pourcentage demandent qu'on soit vraiment familier avec le contenu a évaluer de même que le contenu du test. Et Green termine avec les échelles de nombre de réponses correctes et de niveau d'équivalence qui sont les moins utiles; par contre elles ajoutent des informations aux résultats lorsque employées avec les autres mesures.

L'usage des tests standardisés est souvent accepté dans le monde de l'éducation mais est aussi remis en question. Murphy (1994) mentionne que si les ministères d'éducation des provinces exigent des tests qui couvrent les habiletés de base de même que les connaissances de base dans les sujets majeurs, on peut s'attendre à des résultats de tests superficiels obtenus par les étudiants. On peut se demander pourquoi est-ce que beaucoup de parents, enseignants, groupes ou organismes dans le monde de l'éducation dénigrent, n'utilisent pas ce type de test ou ne croient simplement pas aux résultats de ceux-ci? Jones (1999) fait référence à Sir Lovelace Stamer qui, en 1887,

résume bien l'idée que les personnes se font aux sujets des tests et de leurs résultats. Jones dit que Stamer mentionne qu'un enseignant sait que son statut professionnel dépend sur les résultats académiques que ses élèves obtiennent et qu'il devient une machine pour obtenir de ses élèves de bons résultats; ce qui ne contribue en rien à la cause de l'éducation.

Les avantages et les désavantages des tests standardisés et de leurs résultats sont nombreux et c'est impossible de tous les énumérer. Nous présentons un bref sommaire des avantages et désavantages.

Avantages des tests standardisés et de leurs résultats

Quelques avantages des tests standardisés et de leurs résultats sont: la précision des tests standardisés, leur fidélité et leur validité, leur aisance à tester un grand nombre de personnes en même temps et la vitesse à laquelle ils sont corrigés et analysés. Cizek (1998) mentionne que les résultats de tests standardisés procurent plus d'informations au sujet d'un étudiant à prix modique et ceci dans un court laps de temps. Il mentionne aussi que les maisons de publication et d'analyse de tests standardisés donnent des analyses qui synthétisent et résument bien les informations au sujet des étudiants. En plus, les maisons de publication et d'analyse de tests standardisés utilisent un langage facile à comprendre. Un autre avantage mentionné par Cisek est le niveau élevé de fidélité des tests standardisés qui est de 95% ou plus et que les tests évaluent un vaste éventail d'objectifs incluant les domaines des arts du langage, la mathématique, la science et les habiletés d'étude.

Désavantages des tests standardisés et de leurs résultats

Les désavantages des tests standardisés et de leurs résultats sont: quand les tests sont administrés à un groupe, certaines personnes ne respectent pas les directives du test; l'utilisation partielle et ou erronée des résultats d'un test; l'utilisation de résultats



de tests afin d'évaluer indirectement les professeurs et les écoles pour ne nommer que ceux-ci.

Shepard (1992) explique que les résultats de tests standardisés peuvent mieux servir à identifier des étudiants éligibles à des programmes ou à pointer du doigt les écoles pour les tenir responsables du succès des étudiants. Selon un groupe d'administrateurs d'écoles aux Etats-unis de l'American Association of School Administrators (1991), les données provenants de tests standardisés peuvent être utilisés à mauvais escient. Par exemple certaines écoles établissent des listes d'étudiants qui sont éligibles de fréquenter leur institutions scolaires basées sur des résultats de tests standardisés échelonnés sur une courbe normale et sur la quantité de tests standardisés administrés aux étudiants. Le plus de tests standardisés administrés à un étudiant avec de bons résultats, le plus de chance qu'a un étudiant de fréquenter une institution scolaire.

Mais c'est Cizek (1998) qui résume le mieux les désavantages des tests standardisés. Il discute que le format des tests peut être limité, que les résultats des tests peuvent être énormes en quantité et difficiles à comprendre pour ceux qui n'ont qu'une connaissance limitée des statistiques, que l'enseignement et l'apprentissage peuvent être modifiés pour répondre aux "besoins" du test et à la mauvaise utilisation des tests. Cannell (1990) mentionne l'abus et la mauvaise utilisation de tests standardisés dans la publication du rapport du lac Wobegon. Cannell rapporte comment les enseignants trichaient sur des tests de rendement standardisés administrés dans les écoles. Les enseignants falsifiaient les résultats des tests standardisés, utilisaient des méthodes incorrectes pour tester et possiblement du plagiat.

Comme on peut le constater, les désavantages mentionnés précédemment comme la difficulté d'administrer les tests standardisés et d'en comprendre les résultats, l'utilisation partielle ou erronée de leurs résultats peuvent diminuer l'utilisation des tests standardisés dans le monde de l'éducation et ailleurs. Mais il y a aussi des

avantages comme la fiabilité et la validité des tests standardisés, la facilité à tester un grand nombre de personnes qui font que les tests standardisés sont quand même utilisés dans plusieurs domaines et de façon régulière.

Certaines études mentionnées précédemment sur les tests standardisés nous amène à conclure que la recherche suivante pourra être utile aux enseignants s'il y a un coefficient de corrélation positif significatif entre les résultats du Canadian Achievement Test et la note finale des étudiants. Cette conclusion est basée sur le fait que la majorité des études sur les tests standardisés, sont faites aux Etats-unis et non au Canada et qu'aucune étude n'a été faite sur la corrélation entre les résultats du Canadian Achievement Test et la note finale d'étudiants en neuvième année en anglais et mathématique . Les enseignants pourront ainsi utiliser les résultats de la recherche comme un outil de prédiction de la réussite scolaire des étudiants dans les arts du langage et en mathématique si les corrélations s'avèrent positives et significatives. Les enseignants, en se basant sur les résultats du CAT, pourront prédire avec un certain niveau de précision si un étudiant réussira ou non en anglais et en mathématique.

Description du Canadian Achievement Test

Avant la création du Canadian Achievement Test, certains organismes, écoles et universités au Canada dans les années 1920-30 ont commencé à remettre en question les standards d'évaluation ainsi que les tests utilisés. Tomkins (1986) mentionne certaines inquiétudes énumérées par la conférence nationale des universités canadiennes tenue au début du vingtième siècle. Ces inquiétudes étaient le déclin des standards au secondaire de même que le faible niveau de préparation des étudiants aux examens de fin d'études en vue de l'admission à l'université. Après 1930, l'université de Toronto exige une cinquième année pour l'obtention du diplôme d'études secondaires pour une meilleure préparation des étudiants qui iront à l'université. Suivant l'université de Toronto,

Tomkins (1986) mentionne que plusieurs provinces ont établi une standardisation de leurs programmes d'études et que ceci a amené des standards équivalents entres les provinces. Tomkins mentionne aussi que les universités renforcaient la similitude des standards entre les provinces en exigeant les mêmes standards d'admission pour chaque sujet enseigné au secondaire peu importe la province d'origine de l'étudiant.

En 1936, un comité sous la tutelle de Fred McNally, ministre de l'éducation en Alberta, conclue que la coordination entre les critères d'admission à l'université et les programmes d'études au secondaire était accomplie. Cette standardisation des critères et des programmes incluaient les matières suivantes: littérature et composition anglaise, algèbre et géométrie, chimie et physique, histoire, latin et le français. Mais c'est dans les années 1920 qu'un des premiers programmes scientifiques a été mis sur pied en Colombie-Britannique. Tomkins (1986) mentionne que Peter Sandiford fut engagé par la firme de sondage Putman-Weir pour établir le niveau d'intelligence et d'apprentissage de 17000 étudiants. Putman-Weir croyait que les résultats des tests traditionnels étaient ambigus. Donc des tests furent construits sur les objectifs précis du programme basés sur des méthodes scientifiques et la firme Putman-Weir pensait que leurs résultats seraient plus authentiques que ceux des enseignants et des administrateurs.

Sandiford va donc utiliser des tests américains pour les écoles de la Colombie-Britannique. Il justifie leur utilisation en proclamant que le système éducatif canadien est similaire à celui des Etats-Unis. Cependant la section géographie et histoire est totalement différente donc inutilisable. En conséquence des tests sur ces deux sujets furent construits pour répondre aux exigences spécifiques du programme canadien. L'effet à long terme de la firme Putman-Weir est sans contre-dit la création d'un test national canadien nommé le Dominion Test. Le Dominion Test est le premier test de rendement national basé sur les différents programmes d'études des provinces canadiennes créé par Sandiford. Le Dominion Test est considéré comme étant le précurseur du Canadian Achievement Test. Il faut mentionné toutefois que le Dominion



Test n'était pas considéré comme un test standardisé. Après avoir développé le Dominion Test, Sandiford continua son travail en Ontario et il créa des tests plus fiables que les examens traditionnels créés par les enseignants. En plus ces tests étaient considérés comme ayant de bonnes corrélations avec les notes des écoles (Tomkins, 1986).

Dans les années 1940, Tomkins explique que James Thomson de l'université Queen était contre la standardisation des critères d'admission aux universités canadiennes. Thomson croyait que la diversification des programmes au secondaire était inévitable et que les universités devaient enseigner certaines matières du secondaire pour pallier au lacunes des étudiants dans certaines matières. Il mentionne que l'administration de tests scientifiques complémenterait les examens traditionnels du secondaire et que ces tests aideraient à classifier et catégoriser les lacunes des étudiants du secondaire dans certains sujets.

L'augmentation du nombre d'étudiants allant à l'université dans les années 1950 de même que la diversification des cours offerts au secondaire créa une plus grande différence au niveau du rendement académique des étudiants. Ceci produisit un regain de vie au niveau de l'uniformité des standards dans les écoles canadiennes. Tomkins (1986) mentionne que trois chercheurs de l'Ontario, Ketchum du College Trinity, Smith de l'université de Toronto et Jackson du Ontario College of Education construirent des tests ayant pour but de mesurer d'une manière plus objective le rendement et les habiletés des étudiants. Ils ont voulu se servir de résultats de ces tests pour les admissions à l'université. Ils développèrent le SATO (Scholastic Aptitude Test of Ontario) basé sur trois tests de rendement américains. Donc des tests de rendement provinciaux furent créés pour l'anglais, le français et la chimie. Suivant l 'établissement du SATO, l'étude Carnegie impliqua 120 000 étudiants de la septième à la neuvième année en Ontario pour la mise sur pieds de tests de rendement.

Tomkins cite une étude faite par Manning en 1954 qui résume bien les politiques



des tests au Canada depuis 1945. Manning mentionne que le niveau d'échecs pour les provinces variait d'année en année mais était typiquement élevé pour toutes les provinces notamment en algèbre et en chimie et que conséquemment le haut niveau d'échec était occasionné par un nombre croissant d'étudiants avec des habiletés médiocres ou inférieures. Manning croit qu'en simplifiant les programmes d'études et les exigences des tests de fin du secondaire qu'on pourrait rencontrer d'une manière plus réelle les besoins des étudiants se situant dans la norme.

Dans les années 1960, les universités continuent de déterminer les standards d'admission à l'université. Tomkins (1986) discute que tous les étudiants d'une province doivent écrire les mêmes tests provenant du ministère de l'éducation de chaque province. Fait à noter, ces tests ne diffèrent pas beaucoup d'une province à une autre et la note minimale de passage était la même pour toutes les provinces. Malgré cette uniformité, la demande était forte pour des critères similaires d'admissions aux universités canadiennes. Une des raisons pour vouloir cette unformité était que la Canada était un pays et non un groupe de provinces. Tomkins mentionne qu'un membre du gouvernement de l'Ontario en 1964 du nom de McCarthy croyait que les universités ne pouvaient plus éviter leurs responsabilités vis-à-vis les étudiants et qu'elles devraient augmenter leurs critères d'admission. Donc deux courants distincts apparurent dans les années 1960 qui étaient pour et contre la standardisation des critères d'admission dans les universités canadiennes. Ceci amena en Ontario la naissance d'un organisme, l'OACU (Ontario for Admission to College and University) qui établi un test provincial administré aux élèves de la fin du secondaire comme critère d'admission dans les universités ontariennes. Le programme de l'OACU donna naissance à un programme similaire national institué par le SACU (Service for Admission to College and University) en 1969. Durant cette période, 100 000 étudiants anglophones et francophones du pays écrirent un test standard uniforme comme exigence d'admission à l'université. Les universités ont combiné les résultats de ce test avec le dossier de



l'étudiant du secondaire. Malgré cette initiative d'uniformité, de sérieux problèmes sont apparus. Par exemple le test démontra des lacunes au niveau de la traduction de l'anglais au français. Aussi, plusieurs similarités avec les tests américains étaient présents. Troisièmement parce que le test ne s'éloignait pas suffisamment des programmes d'études de chaque province ce n'était pas possible d'établir des standards similaires canadiens. En dernier lieu, les universités se basaient plus sur les dossiers personnels des étudiants provenants des écoles secondaires que sur les résultats du test Donc en 1972, ce programme national disparut.

Dans les années 1970, Tomkins (1986) mentionne que le monde de l'éducation se concentre sur l'établissement d'un nouveau type d'évaluation avec des standards uniformes pour les provinces canadiennes. Malgré un effort d'uniformisation, les provinces rencontrent des difficultés majeures. Le secteur d'évaluation provincial du Nouveau-Brunswick formé en 1975 développe des outils critériés d'évaluation pour la province. Ces outils critériés sont basés sur des tests créés par la Westinghouse Learning Corporation des Etats-Unis. Ces tests contenaient un biais culturel américain. Une canadianisation du test n'a pas résout le problème des biais culturels.

La Nouvelle-Ecosse établi un programme d'évaluation ayant pour but de complémenter les évaluations scolaires. Ce programme vise à évaluer les cinq sujets majeurs du programme d'éducation de la Nouvelle-Ecosse. Durant cette même période, la Colombie-Britannique développe un programme d'évaluation provincial qui adresse deux questions majeures. Dans quelle mesure les étudiants maîtrisent-ils les objectifs de base des programmes éducatifs de la province et quel pourcentage d'étudiants doit être accepté dans les programmes collégiaux et universitaires? Même si ce programme démontre que les étudiants de la Colombie-Britannique performaient mieux que ceux des Etats-Unis, il indiquait un manque d'habiletés de base lorsque les étudiants entraient sur le marché du travail. L'Alberta opta pour un type d'évaluation semblable à celui de la Colombie-Britannique (Tomkins, 1986).



L'Ontario, cependant, opta pour une différente approche en se basant sur des instruments d'évaluation reliés aux programmes d'études de la province. Tomkins (1986) dit que les instruments de mesure inventés n'étaient pas bien développés pour les niveaux intermédiaires et que leurs effets sur les programmes étaient problématiques. Les résultats de ces tests suggéraient un besoin d'apporter changements au niveau des programmes et du matériel didactique. Puisque ces changements seraient très coûteux pour la province, celle-ci laissa les autorités locales en charge du problème. A la fin des années 1970, les gouvernements provinciaux continuaient à faire face au problème de la standardisation de même qu'a de sérieuses coupures budgétaires. En conséquence les universités ont dû abolir les types d'évaluation provenant de l'extérieur de leurs intitutions. Durant années 1980, l'Alberta développe les tests de rendement pour les niveaux de la troisième, sixième et de la neuvième année (Runté,1998). Pendant cette même période la province réintroduit les examens de diplôme de la douzième année et par les années 1990, l'Alberta va développer des outils diagnostiques en lecture et en mathématiques et établir des indicateurs éducationnels de qualité. Ce qui nous apporte a parler du Canadian Achievement Test.

Le Canadian Achievement Test (1995), communément appelé le CAT, est une batterie de tests qui a été construit en 1981 pour mesurer le rendement des étudiants aux niveaux des habiletés de base en mathématique (au niveau des concepts et du calcul) et dans les arts du langage au niveau lecture, épellation, langage et habiletés d'étude. Le test comprend huit niveaux numérotés de douze à dix-neuf qui correspondent a un niveau de scolarité spécifique. Par exemple, le niveau 18 est pour un niveau de scolarité de 8.0 à 10.2 années (1995) ce qui correspond à un niveau de la huitième à la dixième année. Le test est administré normalement par les enseignants de l'école au début de l'année scolaire, soit au mois de septembre. Le test est écrit en anglais et s'administre sur une période de temps qui totalise 211 minutes et qui est étendue sur une période de cinq jours en moyenne.

Le test se divise en cinq parties: lecture, épellation, langage, habiletés d'étude et mathématique. Les types de mesures utilisées pour chaque partie sont les suivantes: le nombre correct de réponses (NCS), le pointage d'échelle (SS), le rang percentile national (NP), le rang percentile local (LP), l'échelle staneuf nationale (NS), l'échelle staneuf locale (LS), et le niveau d'équivalence (GE). La validité des résultats du CAT est assurée par les différents niveaux du test qui correspondent aux habiletés et concepts du niveau académique des étudiants qui sont normalisés et décrits dans le manuel du Canadian Achievement Test (1995). Les étudiants dans cette étude ont écrit le niveau 18 du test qui correspond à un niveau académique de la huitième année à la dixième année.

Le CAT est administré selon les normes du Class Management Guide (1995). La fidélité des résultats du CAT est mesurée par deux méthodes. La première méthode est la formule 20 Kuder-Richardson ou KR-20. Cette mesure statistique est un coefficient de corrélation basée sur une administration du test à un échantillon donné. La deuxième mesure de la fidélité du test est une mesure d'erreur standard. Cette mesure statistique estime le montant d'erreurs auquel qu'on peut s'attendre dans le résultat du test. Les données décrivant ces caractéristiques statistiques du CAT, rapportées dans le bulletin technique et les suppléments techniques, démontrent que le CAT a une bonne fidélité et validité. Ce test a été créé en se basant sur les différents programmes des provinces canadiennes et à évaluer les habiletés de base au niveau du langage et des mathématiques. Plusieurs revisions furent apportés au test depuis 1981 pour assurer sa validité en fonction des changements des programmes en éducation. Le CAT/2 est une des dernières versions du test à être sur le marché. Voici une description détaillée des cinq types de mesures provenant du Class Management Guide du Canadian Achievement Test.



Le rang percentile national et local

Les rangs percentiles qui ont une valeur de 1 à 99 sont communément utilisés pour informer les parents et les étudiants des résultats du test. Un percentile est interprété comme étant un pourcentage d'étudiants dans un groupe normal qui ont un pointage inférieur au pointage d'un étudiant spécifique. Par exemple, si un étudiant a un rang percentile de 71, cela veut dire que l'étudiant a obtenu un pointage qui serait supérieur à 71% d'un groupe normal d'étudiants. Le rang percentile local compare un étudiant par rapport à un groupe d'étudiant de même niveau dans une école ou conseil scolaire (Canadian Achievement Test 1995).

Les staneufs à l'échelle nationale ou locale

Les staneufs sont des échelles de pointage sur une échelle de neuf unités égales s'étendant de 1 à 9. En général, les staneufs de 1 à 3 sont considérés comme étant des résultats sous la norme, les staneufs de 4 à 6 étant la norme et les staneufs de 7 à 9 comme étant des résultats au-dessus de la norme. Les staneufs sont faciles à comparer et à travailler avec mais sont moins précis que le rang percentile (Canadian Achievement Test 1995).

Le pointage d'échelle

Le pointage d'échelle comprend des sections ou une échelle d'intervalles égaux qui s'applique pour tous les niveaux du CAT indépendamment du niveau scolaire de l'étudiant et de l'année dont le test a été administré. Ces pointages s'étendent sur une échelle de 0 à 999. L'échelle est basée de la même manière qu'on établit un échantillon qui représente la norme ou la moyenne; un pointage de 666 représente la norme ou la moyenne. Lorsque le CAT est administré à un étudiant, cet étudiant devrait obtenir le même pointage d'échelle comparé à la moyenne des autres étudiants et ceci sans tenir compte du niveau ou des méthodes de pointage employées. On se sert aussi du pointage

d'échelle pour calculer le niveau d'équivalence d'un étudiant (Canadian Achievement Test 1995).

Le niveau d'équivalence

Tandis que le rang percentile et les staneufs indiquent un rang de position relatif dans un niveau donné, le niveau d'équivalence s'étend au-delà d'un niveau. L'échelle pour le niveau d'équivalence s'étend de 0.0 à 12.9, ce qui représente les 13 années de scolarité (de la maternelle à la 12 année) et les dix mois de l'année scolaire traditionnelle. Septembre est considéré comme le début de l'année scolaire et est représenté par (.0); octobre est représenté par (.1); etc. Le niveau d'équivalence d'un étudiant représente l'année et le mois scolaire des étudiants faisant partie du groupe moyen et pour qui sa performance au test est théoriquement équivalent au groupe moyen. Pour illustrer le calcul d'un niveau d'équivalence, supposons qu'un test de vocabulaire est administré durant le mois d'avril à un groupe moyen d'étudiants de 6 année et dont la moyenne du pointage d'échelle est de 503. Donc le niveau d'équivalence pour un pointage de 503 pour ce test serait établi à 6.7 le 6 représentant le niveau de 6 ième année et le .7 représentant le mois d'avril. Tout étudiant qui obtiendrait une note de 503 pour ce test de vocabulaire aurait un niveau d'équivalence de 6.7 au niveau du vocabulaire. Il faut être prudent lorsqu'on interprète des niveaux d'équivalences. Par exemple, si un élève de 2 ième année obtient un niveau d'équivalence de 4.8 dans un test de mathématique, cela ne veut pas dire que l'étudiant a maîtrisé tous les concepts enseignés dans ce district scolaire pour les 8 premiers mois de la 4 ième année. Cela veut simplement dire que la performance pour ce test équivaut théoriquement à une performance typique d'un groupe moyen d'étudiants ayant complété huit mois en quatre année (Canadian Achievement Test 1995).



Nombre correct de réponses

Cette dernière échelle du CAT indique tout simplement le nombre de réponses correctes qu'un étudiant a eu soit pour une catégorie spécifique ou pour l'ensemble du test (Canadian Achievement Test 1995).

Question de Recherche

La revue de littérature indique que plusieurs études depuis les années 1900 ont été faites au sujet des tests standardisés. Les résultats de ces études ont démontré que l'utilisation de tests standardardisés peut être utile ou non dans le monde de l'éducation. Ces études ont aussi démontré les avantages et les inconvénients de l'utilisation de tests standardisés. Pratiquement aucune étude n'a été faite avec le Canadian Achievement Test (CAT) qu'on utilise au Canada et particulièrement dans certaines écoles de l'Alberta et rien n'indique qu'on peut utiliser les résultats du Canadian Achievement Test (CAT) comme un indicateur de réussite scolaire des élèves. Aucune étude n'a démontré qu'on peut utiliser les résultats du Canadian Achievement Test (CAT) en relation avec la note finale des étudiants obtenue dans des matières académiques telles que les arts du langage et les mathématiques.

Donc, la questions suivante a guidé cette étude. Est-ce que les résultats de tests standardisés comme le Canadian Achievement Test pourraient être utilisés comme un indicateur du succès académique d'un étudiant en relation avec sa note finale dans un sujet donné?

Les résultats de plusieurs recherches (Hafner 1991, Beard 1990, Johnson 1989, Olson 1989, Baker 1988) ont démontré que l'utilisation de tests standardisés avec d'autres outils d'évaluation peut être utilisée comme un indicateur de réussite des élèves. Pour les fins de cette recherche nous avons tenté d'examiner s'il y a une corrélation positive significative entre les résultats d'une ou plusieurs parties du Canadian

Achievement Test (CAT) et la note finale obtenue par les élèves de niveau de neuvième année dans trois différents programmes: anglais régulier, anglais éducation spéciale et immersion française.

CHAPITRE III

Méthodologie

Premièrement la commission scolaire et l'administration de l'école qui ont participé à l'étude ont été contacté afin d'obtenir leurs permissions pour l'utilisation des résultats. Ce projet a été approuvé par le comité d'éthique de la Faculté Saint-Jean. Le CAT a été administré par les professeurs de la neuvième année de l'école durant la première semaine de septembre 1998 à tous les étudiants de ce niveau. Les données du CAT compilées et analysées par la compagnie PSYCHOMETRICS CANADA LIMITED d'Edmonton, Alberta, Canada ont par la suite été obtenues par le chercheur sur papier et par disquettes informatisées. Les notes finales en mathématique et les arts du langage de tous les étudiants de la neuvième année de l'école en question faisant partie de cette étude ont été obtenues en contactant l'administration de cette école. Ces notes avaient été compilées sur le logiciel SIRS par tous les enseignants en mathématique et les arts du langage pour tous les étudiants de neuvième année de cette école. La note finale des étudiants était composée par la combinaison suivante: tests, quiz, travaux et devoirs basé sur les programmes d'études donnés par les enseignants de l'école qui correspond à 80% de la note finale. La rédaction du test de rendement de la province, écrit en juin correspond à 20% de la note finale pour un total de 100%.

L'obtention des notes finales de l'école de même que les résultats du CAT ont été obtenus des autorités scolaires en juillet 1999. Les données ont été transféré dans le logiciel Excel pour chaque étudiant de neuvième année qui a écrit le CAT et qui a reçu une note finale des enseignants en anglais et en mathématique durant l'année scolaire



1998-99. Chaque partie de l'analyse a reçu un numéro de 1 à 120. Chaque numéro permet ainsi d'organiser les notes finales et les résultats du CAT pour chacun des élèves tout en conservant l'anonymat des étudiants. Les données ont été organisées sous forme de tableau comprenant 7 colonnes. Chaque colonne indique la note finale en math des étudiants, la note finale dans les arts du langage des étudiants, le pointage obtenu dans le CAT pour la lecture, l'épellation, le langage, les habiletés d'étude et la mathématique avec les mesures NCS, SS, GE, NP et NS. Les calculs de corrélations Pearson R ont été faits par les logiciels SPSS et Excel après compilation de toutes les données.

Echantillon

L' échantillon à la base de cette étude se compose d'un total de 120 étudiants de la neuvième année qui ont écrit le Canadian Achievement Test dans une école secondaire premier cycle de la région d'Edmonton et qui ont reçu une note finale en anglais et mathématique de leurs enseignants. Les étudiants qui participent à cette étude proviennent d'un milieu socio-économique aisé.

Les sujets sont répartis dans trois différents programmes; anglais régulier 65%, anglais éducation spéciale15% et immersion française 20%. L'âge des étudiants se situe entre 13 ans et 16 ans au moment d'écrire le CAT. Le pourcentage des étudiants âgés de 13 ans était de 26.7%, 67.5% étaient âgés de 14 ans, 5% étaient âgés de 15 ans et 8.3% étaient âgés de 16 ans. L'échantillon se compose de 47.5% d'étudiants de sexe masculin et de 52.5% d'étudiants de sexe féminin. Un seul étudiant n'a que des résultats partiels pour le CAT test car il a été absent pour raison médicale et les résultats partiels ont été inclus dans l'ensemble des données du CAT car cela n'a aucun effet pour les calculs de corrélation.

Pour des raisons d'éthique, aucun nom d'étudiant n'a été utilisé dans l'étude pour préserver l'anonymat de ceux-ci. Un numéro de 1 à 120 a été attribué à chaque étudiant sur les documents analysés indiquant simplement le fait que les données

réfèrent à un même individu.

Analyse et Interprétation des Résultats

Les résultats de cette recherche sont présentés dans cette partie. Un total de 120 étudiants a participé à cette étude. Les fréquences et les pourcentages de l'échantillon démographique des données sont présentés dans la section échantillon. Les résultats ont été obtenus de l'école qui a participé à l'étude par leurs banques de données. La banque de données incluait le nom, adresse, âge, et genre de tous les étudiants de neuvième année ayant écrit le Canadian Achievement Test, ainsi que les résultats de tous les tests pour chaque individu. Une autre banque de données contenait le nom et la note finale de tous les étudiants de la neuvième année ayant reçu une note finale en anglais et en mathématique.

Comme représenté dans la section échantillon, le pourcentage des étudiants mâles et femelles est presqu'identique. Tous les étudiants proviennent d'un milieu urbain. La majorité d'entre eux (67,5%) sont âgés de 14 ans et sont dans le programme régulier anglais (65%). Un faible pourcentage font partie du programme d'immersion (20%) et d'éducation spéciale (15%).

Les tableaux 1 à 5 représentent les calculs de coefficients de corrélation pour le nombre correct de réponse (NCS), pour le pointage d'échelle (SS), pour le niveau d'équivalence (GE), pour le rang percentile national (NP) et pour l'échelle staneuf nationale (NS) pour toute la population de 120 étudiants. Chaque tableau inclue les sections lecture, épellation, langage, habiletés d'études et mathématique en fonction de la note finale en anglais et en mathématique.

Les tableaux 6 à 20 représentent les calculs de coefficients de corrélation pour le nombre correct de réponse (NCS), pour le pointage d'échelle (SS), pour le niveau d'équivalence (GE), pour le rang percentile national (NP) et pour l'échelle staneuf

nationale (NS) pour les étudiants en anglais, immersion française et éducation spéciale. Chaque tableau inclue les sections lecture, épellation, langage, habiletés d'études et mathématique en fonction de la note finale en anglais et en mathématique.

Le (NCS) représente le nombre correct de réponses qu'un élève a reçu pour une section ou le total du CAT. Le pointage d'échelle (SS) comprend une échelle d'intervalles égaux de 0 à 999 qui s'applique pour tous les niveaux du CAT test indépendamment du niveau scolaire de l'étudiant et de l'année dont le test a été administré. L'échelle est basée de la même manière qu'on établit un échantillon qui représente la norme ou la moyenne; un pointage de 666 représente la norme ou la moyenne. Le pointage d'échelle est aussi utilisé pour calculer le niveau d'équivalence d'un étudiant. Le niveau d'équivalence (GE) indique l'année et le mois scolaire de l'élève faisant partie du groupe moyen et pour qui sa performance au test est théoriquement équivalent au groupe moyen. Le rang percentile national (NP) représente le pourcentage d'un groupe d'étudiants qui ont un pourcentage inférieur comparé à un étudiant spécifique. Et l'échelle staneuf nationale (NS) représente un pointage qui se situe de 1 à 9; 1 étant le pointage le plus bas sous la norme et 9 étant le pointage le plus haut au-dessus de la norme (Canadian Achievement Tests Norm Book 1995).



Tableau 1

Corrélations entre la section NCS (nombre correct de réponses) du CAT et la note finale en anglais et mathématique (Ang, Imm, Es)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.784*	0.585*	0.469*	0.633*	0.593*	0.467*
Note math		-	0.380*	0.337*	0.528*	0.479*	0.613*
Lecture			-	0.602*	0.735*	0.764*	0.562*
Epellation				-	0.756*	0.614*	0.442*
Langage					-	0.807*	0.633*
Etude						-	0.620*
Math							-

* $p<0.01$ (Bilatéral) n=120

Comme on peut le voir dans le tableau 1, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Gay (1981) mentionne que des coefficients de corrélation inférieur à 0,50 ne sont pas significatifs; que des coefficients de corrélation se situant entre 0,60 et 0,80 sont satisfaisants ou significatifs et que des coefficients de corrélation supérieurs à 0,80 sont considérés très significatifs. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,58 et 0,63. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,47 avec les sections NCS (nombre de réponses correctes).

Tableau 2

Corrélations entre la section SS (pointage d'échelle) du CAT et la note finale en anglais et mathématique (Ang, Imm, Es)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.784*	0.580*	0.473*	0.630*	0.594*	0.421*
Note math		-	0.366*	0.338*	0.529*	0.473*	0.506
Lecture			-	0.612*	0.719*	0.756*	0.461*
Epellation				-	0.767*	0.618*	0.353*
Langage					-	0.795*	0.498*
Etude						-	0.471*
Math							-

* p<0.01 (Bilatéral) n=120

Comme on peut le voir dans le tableau 2, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,50 et 0,63. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,47.



Tableau 3

Corrélations entre la section GE (niveau d'équivalence) du CAT et la note finale en anglais et mathématique (Ang, Imm, Es)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.784*	0.585*	0.455*	0.545*	0.583*	0.457*
Note math		-	0.392*	0.332*	0.454*	0.466*	0.599*
Lecture			-	0.625*	0.733*	0.773*	0.585*
Epellation				-	0.747*	0.628*	0.446*
Langage					-	0.798*	0.609*
Etude						-	0.621*
Math							-

* $p<0.01$ (Bilatéral) n=120

Dans le tableau 3, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,55 et 0,60. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,45.

Tableau 4

Corrélations entre la section NP (rang percentile national) du CAT et la note finale en anglais et mathématique (Ang, Imm, Es)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.784*	0.591*	0.462*	0.600*	0.590*	0.491*
Note math		-	0.393*	0.328*	0.503*	0.475*	0.618*
Lecture			-	0.619*	0.727*	0.773*	0.593*
Epellation				-	0.761*	0.630*	0.460*
Langage					-	0.814*	0.657*
Etude						-	0.641*
Math							-

* $p<0.01$ (Bilatéral) n=120

Dans le tableau 4, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,59 et 0,62. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,46.



Tableau 5

Corrélations entre la section NS (échelle staneuf national) du CAT et la note finale en anglais et mathématique (Ang, Imm, Es)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.784*	0.564*	0.482*	0.637*	0.569*	0.484*
Note math		-	0.349*	0.328*	0.537*	0.449*	0.619*
Lecture			-	0.611*	.0705*	0.7298	0.527*
Epellation				-	0.738*	0.590*	0.414*
Langage					-	0.756*	0.616*
Etude						-	0.567*
Math							-

* p<0.01 (Bilatéral)   n=120

Dans le tableau 5, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,56 et 0,64. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,48.



Tableau 6

Corrélations entre la section NCS (nombre correct de réponses) du CAT et la note finale en anglais et mathématique (Anglais)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.779*	0.589*	0.354*	0.594*	0.542*	0.436*
Note math		-	0.388*	0.289	0.505*	0.428*	0.595*
Lecture			-	0.430*	0.623*	0.715*	0.494*
Epellation				-	0.674*	0.500*	0.404*
Langage					-	0.799*	0.585*
Etude						-	0.541*
Math							-

* p<0.01 (Bilatéral)   n=78

Dans le tableau 6, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,54 et 0,59. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,35.



Tableau 7

Corrélations entre la section SS (pointage d'échelle) du CAT et la note finale en anglais et mathématique (Anglais)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.779*	0.570*	0.368*	0.581*	0.543*	0.437*
Note math		-	0.363*	0.297	0.499*	0.421*	0.593*
Lecture			-	0.477*	0.603*	0.707*	0.496*
Epellation				-	0.690*	0.517*	0.417*
Langage					-	0.791*	0.578*
Etude						-	0.532*
Math							-

* p<0.01 (Bilatéral) n=78

Dans le tableau 7, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,54 et 0,59. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,36.



Tableau 8

Corrélations entre la section GE (niveau d'équivalence) du CAT et la note finale en anglais et mathématique (Anglais)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.774*	0.599*	0.355*	0.505*	0.548*	0.416*
Note math		-	0.409*	0.291*	0.427*	0..427*	0.576*
Lecture			-	0.474*	0.590*	0.716*	0.519*
Epellation				-	0.671*	0.538*	0.404*
Langage					-	0.790*	0.537*
Etude						-	0.523*
Math							-

* p<0.01 (Bilatéral) n=78

Dans le tableau 8, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,50 et 0,59. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,35.



Tableau 9

Corrélations entre la section NP (rang percentile national) du CAT et la note finale en anglais et mathématique (Anglais)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.779*	0.617*	0.357*	0.566*	0.545*	0.464*
Note math		-	0.423*	0.289*	0.490*	0.435*	0.603*
Lecture			-	0.462*	0.589*	0.727*	0.519*
Epellation				-	0.681*	0.518*	0.427*
Langage					-	0.798*	0.607*
Etude						-	0.565*
Math							-

* p<0.01 (Bilatéral) n=78

Dans le tableau 9, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,54 et 0,61. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,35.



Tableau 10

Corrélations entre la section NS (échelle staneuf national) du CAT et la note finale en anglais et mathématique (Anglais)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.774*	0.559*	0.380*	0.587*	0.504*	0.463*
Note math		-	0.345*	0.297	0.490*	0.385*	0.607*
Lecture			-	0.437*	0.615*	0.703*	0.478*
Epellation				-	0.657*	0.482*	0.394*
Langage					-	0.764*	0.578*
Etude						-	0.501*
Math							-

* $p<0.01$ (Bilatéral) n=78

Dans le tableau 10, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de la lecture et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de l'épellation. Les sections lecture, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,50 et 0,60. La section épellation est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,38.



Tableau 11

Corrélations entre la section NCS (nombre correct de réponses) du CAT et la note finale en anglais et mathématique (Education spéciale)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.253	0.426*	0.268	0.686*	0.046	- 0.049
Note math		-	- 0.370	- 0.171	0.211	- 0.136	0.468*
Lecture			-	0.379*	0.598*	0.462*	- 0.100
Epellation				-	0.559*	0.000	- 0.370
Langage					-	0.105	0.038
Etude						-	0.369*
Math							-

* $p<0.01$ (Bilatéral) n=18

Dans le tableau 11, le coefficient de corrélation positif le plus fort se situe seulement au niveau du langage avec un coefficient de corrélation de 0,68. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation, des habiletés d'étude et de la mathématique.. Les sections lecture, épellation, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs non significatifs avec des coefficients de corrélation se situant entre 0,26 et 0,46.

Tableau 12

Corrélations entre la section SS (pointage d'échelle) du CAT et la note finale en anglais et mathématique (Education spéciale)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.305*	0.415*	0.251	0.683*	- 0.009	- 0.117
Note math		-	- 0.379	- 0.194	0.245	- 0.012	0.544*
Lecture			-	0.348*	0.575*	0.410*	- 0.209
Epellation				-	0.538*	- 0.008	- 0.379
Langage					-	0.106	0.001
Etude						-	0.292
Math							-

* p<0.01 (Bilatéral) n=18

Dans le tableau 12, les coefficients de corrélation positifs les plus forts se situent au niveau du langage et de la mathématique avec un coefficient de corrélation de 0,68 et 0,54. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation, des habiletés d'étude. Les sections lecture et épellation ont des coefficients de corrélation positifs non significatifs de 0,25 et 0,41. La section habiletés d'étude a un coefficient de corrélation négatif non significatif de -0,01.



Tableau 13

Corrélations entre la section GE (niveau d'équivalence) du CAT et la note finale en anglais et mathématique (Education spéciale)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.305*	0.420*	0.252	0.624*	- 0.039	- 0.128
Note math		-	- 0.357	- 0.176	0.185	- 0.075	0.532*
Lecture			-	0.319*	0.626*	0.447*	- 0.214
Epellation				-	0.575*	- 0.047	- 0.385
Langage					-	0.168	- 0.037
Etude						-	0.277
Math							-

* p<0.01 (Bilatéral) n=18

Dans le tableau 13, les coefficients de corrélation positifs les plus forts se situent au niveau du langage et de la mathématique avec un coefficient de corrélation de 0,62 et 0,53. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation, des habiletés d'étude. Les sections lecture et épellation ont des coefficients de corrélation positifs non significatifs de 0,41 et 0,25. La section habiletés d'étude a un coefficient de corrélation négatif non significatif de -0,03.

Tableau 14

Corrélations entre la section NP (rang percentile national) du CAT et la note finale en anglais et mathématique (Education spéciale)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.305*	0.411*	0.313*	0.591*	- 0.018	- 0.067
Note math		-	- 0.325	- 0.132	0.130	- 0.082	0.544*
Lecture			-	0.394*	0.672*	0.438*	- 0.043
Epellation				-	0.671*	0.015	- 0.385
Langage					-	0.185	0.032
Etude						-	0.308*
Math							-

* p<0.01 (Bilatéral) n=18

Dans le tableau 14, les coefficients de corrélation positifs les plus forts se situent au niveau du langage et de la mathématique avec un coefficient de corrélation de 0,59 et 0,55. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation, des habiletés d'étude. Les sections lecture et épellation ont des coefficients de corrélation positifs non significatifs de 0,41 et 0,31. La section habiletés d'étude a un coefficient de corrélation négatif non significatif de -0,01.

Tableau 15

Corrélations entre la section NS (échelle staneuf national) du CAT et la note finale en anglais et mathématique (Education spéciale)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.305*	0.378*	0.325*	0.663*	- 0.098	- 0.080
Note math		-	- 0.358	- 0.259	0.286	- 0.104	0.606*
Lecture			-	0.423*	0.465*	0.493*	- 0.254
Epellation				-	0.605*	- 0.062	- 0.463
Langage					-	- 0.111	- 0.092
Etude						-	0.181
Math							-

* p<0.01 (Bilatéral)   n=18

Dans le tableau 15, les coefficients de corrélation positifs les plus forts se situent au niveau du langage et de la mathématique avec un coefficient de corrélation de 0,66 et 0,60. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation, des habiletés d'étude. Les sections lecture et épellation ont des coefficients de corrélation positifs non significatifs de 0,37 et 0,32. La section habiletés d'étude a un coefficient de corrélation négatif non significatif de -0,09.

Tableau 16

Corrélations entre la section NCS (nombre correct de réponses) du CAT et la note finale en anglais et mathématique (Immersion française)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.890*	0.400*	0.543*	0.657*	0.582*	0.558*
Note math		-	0.190	0.313*	0.557*	0.569*	0.723*
Lecture			-	0.670*	0.597*	0.510*	0.467*
Epellation				-	0.797*	0.580*	0.130
Langage					-	0.699*	0.347*
Etude						-	0.302*
Math							-

* p<0.01 (Bilatéral) n=24

Dans le tableau 16, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de l'épellation et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de la lecture. Les sections épellation, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,54 et 0,72. La section lecture est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,39.



Tableau 17

Corrélations entre la section SS (pointage d'échelle) du CAT et la note finale en anglais et mathématique (Immersion française)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.890*	0.425*	0.538*	0.666*	0.581*	0.566*
Note math		-	0.217	0.314*	0.569*	0.566*	0.733*
Lecture			-	0.660*	0.594*	0.508*	0.050
Epellation				-	0.794*	0.564*	0.123
Langage					-	0.671*	0.366*
Etude						-	0.298
Math							-

* p<0.01 (Bilatéral)   n=24

Dans le tableau 17, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études, de l'épellation et de la mathématique. Le coefficient de corrélation le plus bas se situe au niveau de la lecture. Les sections épellation, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,53 et 0,73. La section lecture est la seule ayant un coefficient de corrélation positif non significatif avec une valeur de 0,42.

Tableau 18

Corrélations entre la section GE (niveau d'équivalence) du CAT et la note finale en anglais et mathématique (Immersion française)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.890*	- 0.066	0.481*	0.155	0.571*	0.564*
Note math		-	- 0.278	0.260	0.283	0.567*	0.735*
Lecture			-	0.186	- 0.009	- 0.041	- 0.262
Epellation				-	0.068	0.644*	0.111
Langage					-	0.230	0.266
Etude						-	0.264
Math							-

* $p<0.01$ (Bilatéral) n=24

Dans le tableau 18, les coefficients de corrélation positifs les plus forts se situent au niveau des habiletés d'études et de la mathématique avec des coefficients de corrélation de 0,57 et 0,73. Les coefficients de corrélation les plus bas se situent au niveau de la lecture, de l'épellation et du langage. Les sections épellation et langage ont des coefficients de corrélation positifs non significatifs de 0,48 et 0,15. La section lecture est la seule ayant un coefficient de corrélation négatif non significatif avec une valeur de -0,06.

Tableau 19

Corrélations entre la section NP (rang percentile national) du CAT et la note finale en anglais et mathématique (Immersion française)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.890*	0.379*	0.492*	0.612*	0.572*	0.566*
Note math		-	0.175	0.260	0.512*	0.564*	0.734*
Lecture			-	0.689*	0.614*	0.488*	0.028
Epellation				-	0.784*	0.618*	0.087
Langage					-	0.674*	0.289
Etude						-	0.242
Math							-

* p<0.01 (Bilatéral) n=24

Dans le tableau 19, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études et de la mathématique. Les coefficients de corrélation les plus bas se situent au niveau de la lecture et de l'épellation. Les sections, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,57 et 0,73. Les sections lecture et épellation ont des coefficients de corrélation positif non significatifs de 0,37 et de 0,49.

Tableau 20

Corrélations entre la section NS (échelle staneuf national) du CAT et la note finale en anglais et mathématique (Immersion française)

	Note anglais	Note math	Lecture	Epellation	Langage	Etude	Math
Note anglais	-	0.890*	0.421*	0.499*	0.685*	0.531*	0.541*
Note math		-	0.191	0.252	0.619*	0.518*	0.667*
Lecture			-	0.628*	0.504*	0.429*	- 0.006
Epellation				-	0.706*	0.551*	0.029
Langage					-	0.660*	0.379*
Etude						-	0.194
Math							-

* p<0.01 (Bilatéral) n=24

Dans le tableau 20, les coefficients de corrélation positifs les plus forts se situent au niveau du langage, des habiletés d'études et de la mathématique. Les coefficients de corrélation les plus bas se situent au niveau de la lecture et de l'épellation. Les sections, langage, habiletés d'étude et mathématique ont tous des coefficients de corrélation positifs significatifs avec des coefficients de corrélation se situant entre 0,53 et 0,68. Les sections lecture et épellation ont des coefficients de corrélation positif non significatifs de 0,42 et de 0,49.

Les résultats des tableaux un à cinq démontrent que plusieurs coefficients de corrélations sont adéquats ou significatifs pour la population totale incluant les trois programmes. Les coefficients de corrélation pour le langage sont tous de 0,60 ou supérieur sauf pour la mesure GE donc significatifs. Les résultats pour la mathématique indiquent aussi des coefficients de corrélation près de 0,60 donc significatifs sauf pour la

mesure SS. Les coefficients de corrélation les plus bas se situent tous au niveau de l'épellation avec des mesures inférieure à 0,50 donc non significatifs. Donc, les sections langage et mathématique avec des coefficients de corrélations d'environ 0,60 ou plus pourraient être utilisés comme des indicateurs de réussite scolaire de l'élève.

Les résultats des tableaux six à dix démontrent que plusieurs coefficients de corrélations sont adéquats ou significatifs pour la population étudiante faisant partie du programme d'anglais régulier. La section mathématique pour les mesures NP, NCS, SS et NS et la section lecture pour la mesure NP et GE ont des coefficients de corrélation d'environ 0,60 ou plus. Donc on pourrait les utiliser comme des indicateurs de réussite scolaire de l'élève. Les coefficients de corrélations les plus bas se situent tous au niveau de l'épellation pour les cinq mesures avec des coefficients de corrélation non significatifs inférieurs à 0,50.

Les résultats des tableaux 11 à 15 démontrent que quelques coefficients de corrélations sont adéquats ou significatifs pour la population étudiante faisant partie du programme d'éducation spéciale. La section langage pour les cinq mesures et la section mathématique pour la mesure NS ont des coefficients de corrélation d'environ 0,60 ou plus. Donc on pourrait seulement utiliser la section langage comme un indicateur de réussite scolaire de l'élève. Tous les autres coefficients de corrélations sont non significatifs inférieurs à 0,50. Quelques coefficients de corrélations sont mêmes négatifs.

Les résultats des tableaux 16 à 20 démontrent que plusieurs coefficients de corrélations sont adéquats ou significatifs pour la population étudiante faisant partie du programme d'immersion française. La section mathématique pour les cinq mesures et la section langage pour les mesures NCS, SS, NP et NS ont des coefficients de corrélation d'environ 0,60 ou plus. Donc on pourrait les utiliser comme des indicateurs de réussite scolaire de l'élève. Les coefficients de corrélations les plus bas se situent tous au niveau de la lecture pour les cinq mesures avec des coefficients de corrélation non significatifs inférieurs à 0,50. A noter que la section mathématique pour l'immersion française a reçu

les coefficients de corrélation positifs les plus élevés avec des mesures se situant entre 0,66 et 0,73.

CHAPITRE IV

Discussion

A partir des résultats des tableaux un à vingt, nous pouvons conclure qu'il existe des coefficients de corrélation significatifs entre plusieurs sections du Canadian Achievement Test et la note finale des étudiants dans les arts du langage (anglais) et la mathématique.

Les sections lecture, langage, habiletés d'étude et mathématique des cinq premiers tableaux ont presqu'exclusivement des coefficients de corrélation qui se situent entre 0,57 et 0,64 ceci pour toute l'échantillon de 120 étudiants répartis dans les trois programmes incluant les cinq mesures utilisées dans le CAT c'est-à-dire le nombre correct de réponse (NCS), le pointage d'échelle (SS), le niveau d'équivalence (GE), le rang percentile national (NP) et l'échelle staneuf nationale (NS). Ces coefficients de corrélation sont donc significatifs ou très près du niveau d'être significatif qui est de 0,60 (Gay 1981). Les seules exceptions dans les sections mentionnées ci-haut se situent dans le pointage d'échelle en mathématique et en habiletés d'étude avec un coefficient de corrélation de 0,47 et de 0,50 respectivement. La section épellation par contre n'a obtenu aucun coefficient de corrélation significatif en fonction de la note finale des étudiants. Tous les coefficients de corrélation pour les cinq mesures utilisées dans le CAT se situent entre 0,45 et 0,48 qui se situe au dessous de 0,50 donc qui ne sont pas significatifs.

Les tableaux six à dix représentent 78 étudiants de l'échantillon totale de 120 étudiants qui sont dans le programme d'anglais régulier. La section mathématique pour les mesures NP, NCS, SS et NS et la section lecture pour la mesure NP et GE ont des

coefficients de corrélation d'environ 0,60 ou plus. Ces coefficients de corrélation sont donc significatifs ou très près du niveau d'être significatif qui est de 0,60 (Gay 1981). Les coefficients de corrélation les plus bas se situent tous au niveau de l'épellation pour les cinq mesures avec des coefficients de corrélation non significatifs inférieurs à 0,50.

Les tableaux 11 à 15 représentent 18 étudiants de l'échantillon totale de 120 étudiants qui sont dans le programme d'éducation spéciale. Seule la section langage pour les cinq mesure et la section mathématique pour la mesure NS démontrent des coefficients positifs de corrélation de 0,60 ou plus. Ces coefficients de corrélation sont donc significatifs ou très près du niveau d'être significatif qui est de 0,60 (Gay 1981). Tous les autres coefficients de corrélation sont inférieurs à 0,50 donc non significatifs.

Et les tableaux 16 à 20 représentent 24 étudiants de l'échantillon totale de 120 étudiants qui sont dans le programme d'immersion française. La section mathématique pour les cinq mesures et la section langage pour les mesures NCS, SS, NP et NS ont des coefficients de corrélation d'environ 0,60 ou plus. Ces coefficients de corrélation sont donc significatifs ou très près du niveau d'être significatif qui est de 0,60 (Gay 1981). Les coefficients de corrélation les plus bas se situent tous au niveau de la lecture pour les cinq mesures avec des coefficients de corrélation non significatifs inférieurs à 0,50.

D'une manière générale, nous pouvons donc déduire en se basant sur notre hypothèse de départ qui était de vérifier s'il y avait une corrélation positive significative entre les résultats du Canadian Achievement Test et la note finale en mathématique et en anglais que l'on peut se servir de plusieurs sections du test comme un indicateur de prédiction de la réussite scolaire d'un étudiant. Nos résultats démontrent que plusieurs corrélations étaient significatives sauf pour l'éducation spéciale. Les résultats vont dans une direction similaire de plusieurs recherches faites dans des domaines similaires depuis une quinzaine d'années environ.

Sue et Abe (1984) ont comparé la moyenne académique d'étudiants en fonction de résultats de tests standardisés et ont obtenu des résultats qui pouvaient prédire la

note d'étudiants à l'université de facon modérée. Leur échantillon comprenait plus de 5000 étudiants ce qui était très supérieur à notre échantillon de recherche. Pratt (1987) a obtenu des corrélations très fortes entre les résultats d'un test standardisé, le National Teachers Examination et la note finale à l'université. Son échantillon ne comprenait que 64 étudiants. Notre échantillon de recherche est le double de celui de Pratt. Pearlman (1988) a obtenu aussi des corrélations fortes entre les résultats de deux tests standardisés. Son échantillon était de 400 étudiants. Johnson (1989) a obtenu des résultats qui pointent dans la même direction que notre recherche. Johnson (1989) démontre qu'il y avait un indicateur significatif entre des variables en lecture et la note reçue en première année d'université. En dernier lieu, Hafner (1991) a calculé des coefficients de corrélation entre plusieurs tests standardisés et a obtenu une corrélation de 0,77 pour la compréhension en lecture et pour le vocabulaire et une corrélation de 0,88 en mathématique.

Les résultats bas pour l'épellation sont peut-être reliés à différents facteurs. Premièrement, les étudiants en immersion ont reçu la majorité de leur éducation en français et le CAT est un test de langue anglaise. Deuxièmement, les étudiants en éducation spéciale ne suivent pas le même programme d'étude ce qui peut aussi affecter les résultats. En dernier lieu, l'épellation au secondaire est peut-être moins enseigné qu'à l'élémentaire ce qui peut affecter les résultats. Par contre la mathématique et le langage ont des résultats plus élevés. Pour la mathématique, cela est peut-être occasionné par la séparation de deux parties seulement dans le test (concepts et application des concepts et habiletés de calcul) tandis que la section anglais a cinq parties. Pour le langage, les résultats sont peut-être affectés du fait que les étudiants du programme d'immersion sont bilingues et ont peut-être des habiletés supérieures à la norme.

Une des limitation de l'étude est que l'échantillon au complet ne provient que d'une seule école. Alors les résultats ne peuvent pas être généraliser à toutes les écoles de secondaire premier cycle ayant écrit le Canadian Achievement Test. Une deuxième

limitation est la taille de l'échantillon. L'échantillon se compose seulement de 120 étudiants donc on ne peut pas généraliser les résultats à toute la population de la province ayant écrit le CAT et se situant dans le même groupe d'âge mais on peut les transférer pour les fins d'autres études ou recherches. Cette limitation s'applique encore plus lorsqu'on divise la population en trois catégories soit anglais régulier, éducation spéciale et immersion française.

Une troisième limitation est le nombre restreint des étudiants provenant des programmes d'immersion française (24), d'anglais éducation spéciale (18) en comparaison avec le nombre d'étudiants du programme régulier anglais (78). Un nombre plus proportionnel d'étudiants entre les programmes d'immersion française et d'éducation spéciale et le programme d'anglais régulier serait peut-être plus favorable car le nombre d'étudiants des différents programmes de la population serait plus égal.

Une autre limitation est que cette recherche a utilisé une population provenant d'un milieu socio-économique aisé et que la majorité des étudiants proviennent de la même race (caucasienne).

Une autre limitation est qu'en ayant une valeur diagnostique qui indique où se situe certains problèmes, les enseignants et les administrateurs devront trouver des solutions pour résoudres ces problèmes et peut`-être qu'économiquement, cela peut causer des difficultés.

Et en dernier lieu, peut-être qu'en déterminant en avance si un étudiant est plus à risque d'échouer ou de passer que ceci peut indirectement influencé certains professeurs et administrateurs en créant certaines attentes basées sur ces résultats.

Malgré ces limitations, les résultats de cette recherche pourront être utilisés dans le futur par des écoles ou districts scolaires utilisant le CAT et ayant un ou des individus ayant les mêmes caractéristiques que l'échantillon utilisé dans cette recherche. Ils pourront utiliser les résultats de cette recherche comme un outil diagnostique pour prédire le succès scolaire des étudiants. Les hypothèses générées par cette étude

peuvent être vérifier par d'autres chercheurs dans des milieux à caractères similaires.

Conclusion

Le but de notre étude était de démontrer s'il y avait une corrélation significative entre les résultats du Canadian Achievement Test et la note finale obtenue en anglais et en mathématique par des étudiants de neuvième année. Les sections lecture, langage, habiletés d'étude et mathématique du CAT ont obtenues des coefficients de corrélation significatifs en fonction de la note finale obtenue par des étudiants en anglais et en mathématique. Ce qui veut dire qu'on peut utiliser ces parties du CAT comme des indicateurs de réussite scolaire en anglais et en mathématique. Par contre les résultats obtenus en épellation ne sont aucunement significatif et ne peuvent pas être utilisés comme un indicateur de réussite scolaire.

Voici quelques suggestions qui pourront être utiles pour des recherches similaires à l'avenir. Premièrement, toute recherche qui utilisera totalement ou en partie le Canadian Achievement Test est fortement recommendée car l'utilisation du CAT est répandue au Canada. Cette recherche pourra aidé toutes les écoles ou districts scolaires qui administrent le CAT test car ils pourront utiliser les résultats du test comme un outil diagnostique. Nous recommendons une recherche similaire avec un échantillon plus ample ce qui pourrait apporter une généralisation plus fiable pour l'ensemble de la province ou du pays. Et cette étude peut servir comme base pour une étude plus elaborée incluant toutes les provinces du Canada. Nous considérons qu'un échantillon de plus que 120 étudiants serait bénéfique pour ce type de recherche et que l'échantillon soit tiré dans plus qu'une école qui représenterait plusieurs niveaux socio-économiques. Une recherche similaire représentant seulement des programmes d'immersion française ou francophone serait utile étant donné le peu de recherches faites dans ces deux programmes en Alberta.

Cette recherche peut aussi remettre en question l'utilisation du CAT test non

seulement comme un test standardisé mais aussi comme un outil diagnostique qui peut prédire la réussite scolaire d'un étudiant en anglais et en mathématique. Cette recherche peut aussi apporter des questions sur l'utilisation des tests standardisés en milieu scolaire qui ne sont pas mandatés par la majorité des provinces canadiennes. Pourquoi est-ce que le CAT est utilisé dans les écoles? Pourquoi est-ce que le CAT est administré au début de l'année scolaire par l'école qui a participé à cette étude? Est-ce que les deux mois de vacances ont une influence sur les résultats du CAT? Est-ce que les résultats de cette recherche serait différents si le CAT serait administré à la fin de l'année scolaire donc plus près de l'échéance de la note finale de l'étudiant et non dix mois précédants la note finale? En dernier lieu, nous suggérons des recherches de type similaire mais incluant différents tests standardisés ce qui pourrait aider à tirer des parallèles entre cette recherche et d'autres recherches.

Bibliographie

Angoff, W. H. (1984). Scales, norms, and equivalent scores. New Jersey: Educational Testing Service.

Ascher, C. (1990). Assessing bilingual students for placement and instruction. New York, NY: Clearinghouse on Urban Education.

Baker, E. L. (1988). Mandated tests: reform or quality indicator. Los Angeles, California: Center for Research on Evaluation, Standards, and Student Testing.

Beard, J. G. (1990). The prediction of state student assessment test scores from scores on the districts' standardized norm-referenced tests. Tallahassee, FL: Florida State Department of Education.

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self-concepts. Journal of Educational Psychology, 90 (1), 102-110.

Bunyan, P. (1998). Comparing pupil performance in key stages 2 and 3 science SAT's. School Science Review, 79 (289), 85-87.

Burns, M. (1984). What do test scores really mean? (Rev. ed.). California: Los Angeles Unified School District.

Canadian Test Center, Educational Assessment Services. (1995). Canadian achievement tests, class management guide, Form A. Toronto: CTC/McGraw-Hill Ryerson.

Canadian Education Statistics Council. (1994). A survey of student assessment/evaluation in Canada. Edmonton, Alberta: Author.

Cannell, J. J. (1989). How public educators cheat on standardized achievement tests: The "Lake Wobegon" report. New Mexico: Friends for Education. (ERIC Document Reproduction Service No. ED 314 454).

Cisek, G. (1998). Filling in the blanks: Putting standardized tests to the test. Fordham Report, 2 (11), 64.

Creech, J. D. (1998). How do SREB states gauge student achievement? Atlanta, GA: Southern Regional Education Board. (ERIC Document Reproduction Service No. Ed 426 505).

Demaray, M. K., & Elliott, S. N. (1998). Teachers' judgments of students' academic functioning: A comparison of actual and predicted performances. School Psychology Quarterly, 13 (1), 8-24.

Ebel, R. L., & Frisbie, D. (1986). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall.

Fenton, R. (1998). Anchorage school district profile of performance 1997-1998. assessment and evaluation report (Rep. No. 98-3). Anchorage, AK: Anchorage School District. (ERIC Document Reproduction Service No. ED 427 086).

Gay, L. R. (1981). Educational research: competencies for analysis and application. Columbus, OH: Merrill.

Green, D. R. (1987). Interpreting scores from standardized achievement tests. National Association of Secondary Schools Bulletin, 71, (496), 23-35.

Gronna, S. S., Jenkins, A., & Chin-Chance, S. (1998). The performance of students with disabilities in a norm-referenced, statewide standardized testing program. Journal of Learning Disabilities, 31, (5), 482-493.

Hafner, L., & Stakenas, R. (1990). A study of the social reading achievement and reading interests of a group of eighth graders. Florida: Florida State University Developmental Research School. (ERIC Document Reproduction Service No. Ed 317 957).

Henriksson, W., & Wolming, S. (1998). Academic performance in four study programmes: A comparison of students admitted on the basis of GPA and SweSAT scores, with and without credits for work experience. Scandinavian Journal of Educational Research, 42, (2), 135-150.

House, D. J., & Prion S. K. (1998). Students attitudes and academic background as predictors of achievement in college English. International Journal of Instructional Media, 25, (1), 29-42.

Impara, J., & Plake, B. (Eds.). (1998). The thirteenth mental measurements yearbook. Lincoln, NB: University of Nebraska Press.

Johnson, L. R. (1989, March). Ten percent more: What learning process dimensions can add to the prediction of first year college students' academic success. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Jones, L. V. (1999, April). The assessment of student achievement: The hundred years war. Paper presented at the Annual Meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Linn, R. L. (1991). Test misuse: Why is it so prevalent? Colorado: Congress of the US, Washington, DC. Office of Technology Assessment.

Mayer, D. P. (1998). Do new teaching standards undermine performance on old tests? Educational Evaluation & Policy Analysis, 20, 53-73.

McCormick, A. C. (1999). Credit production and progress toward the bachelor's degree: An analysis of postsecondary transcripts for beginning students at 4-year institutions. Educational Statistics Quarterly, 1, (1), 75-76.

McCurdy, J., & Speich, D (Eds.). (1991). Testing (Special issue). Education USA, 33, (19).

Menchaca, V. D. (1992, January). Achievement motivation in eighth grade students of two ethnic groups. Paper presented at the Meeting of the Southwest Educational Research Association, Houston, TX.

Mehrens, W. A., & Lehmann, I. J. (1991). Measurement and evaluation in education and psychology. Fort Worth: Holt Rinehart & Winston.

Miller, R. H., & Pennington, S. C. (1988, October). Teachers as researchers: Does the DRP predict student achievement? Paper presented at the Annual Meeting of the Florida Reading Association, Orlando, Fl.

Murphy, S. (1994). "No-one has ever grown taller as a result of being measured"-Six educational measurement lessons for Canadians. In L. Erwin & D. MacLennan (Eds.), Sociology of education in Canada: Critical perspectives on theory, research & practice (pp. 238-252). Toronto, Ont.: Copp Clark Longman.

Nelson, R. (1998). Using a student performance framework to analyze success and failure. Journal of College Reading & Learning, 29, (1), 82-89.

Olson, G. H. (1989, March). On the validity of performance grades: The relationship between teacher-assigned grades and standard measures of subject matter acquisition. Paper presented at the Annual Meeting of the National Council on Measurement in Education, San Francisco, CA.

Perlman, C. L. (1989, March). The IGAP and the ITBS: A comparative study. Paper presented at the Annual Meeting of the National Council of Measurement in Education, San Francisco, CA.

Perrone, V. (1991). Position paper: On standardized testing. Childhood Education, 67, 132-142.

Pratt, L. (1987, October). Predicting student performance on the professional knowledge portion of the NTE core battery. Paper presented at the Annual Conference of the Southern Association for Institutional Research and the Society for College and University Planning, New Orleans, LA.

Reis, S. M., Westberg, K., Kulikowich, J., & Purcell, J. (1998). Curriculum compacting and achievement test scores: What does the research say? Gifted Child Quarterly, 25, (2), 123-129.

Rebuffot, J. 1993. Le point sur l'immersion au Canada. Anjou, Québec: Centre Educatif et Culturel Inc.

Rosa, C. M. (1989, March). A comparison of two standardized reading and mathematics achievement tests in the native language for Hispanic limited-English-proficient students. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Runté, R. (1998). The impact of centralized examinations on teacher professionalism. Canadian Journal of Education, 23, (2), 166-181.

Shepard, L. A. (1992, April). Chapter 1's part in the juggernaut of standardized testing. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Smith, M. S., Stevenson, D. L., & Li, C. P. (1998). The voluntary national tests: Mobilizing America to improve education. Curriculum Administrator, 33, (7), 36-37.

Sue, S., & Abe, J. (1988). Predictors of academic achievement among Asian American and White students (College Board Rep. No. 88-11). New York, NY: College Entrance Examination Board.

Wakeford, M. E. (1990, March). The use of NTE communication skills and general knowledge scores for admission to teacher education programs: Policy recommendations. Paper presented at the Annual Meeting of the North Carolina Association for Research in Education, Research Triangle Park, NC.

